

CLAIMS

1 1. A method for processing a document that includes a
2 plurality of fields having respective contents, the
3 method comprising:

4 providing labels to be assigned respectively to the
5 fields and one or more rules applicable to the contents
6 of the fields according to the labels assigned thereto;

7 reading the contents of the fields; and

8 assigning the labels to the fields responsive to
9 application of the rules to the contents.

1 2. A method according to claim 1, wherein the rules
2 comprise a criterion to be applied to the contents of a
3 single one of the fields.

1 3. A method according to claim 1, wherein the rules
2 comprise a relation between the contents of two or more
3 of the fields.

1 4. A method according to claim 3, wherein the contents
2 comprise numbers, and wherein the relation comprises a
3 mathematical relationship between the numbers contained
4 in the two or more of the fields.

1 5. A method according to claim 3, wherein the contents
2 comprise alphanumeric characters, and wherein the
3 relation comprises a semantic relationship between words
4 formed by the characters.

1 6. A method according to claim 1, and comprising
2 providing one or more geometrical rules indicating an
3 expected geometrical relationship between two or more of
4 the fields according to the labels assigned thereto, and
5 wherein assigning the labels to the fields comprises

6 applying the geometrical rules along with the rules
7 applicable to the contents of the fields.

1 7. A method according to claim 1, wherein assigning the
2 labels comprises making a test assignment of the labels
3 to the fields, and rejecting the test assignment if the
4 contents of the assigned fields do not satisfy the rules
5 applicable according to the labels.

1 8. A method according to claim 7, wherein making the
2 test assignment comprises iteratively testing
3 substantially all possible assignments of the labels to
4 the fields, so as to find one or more candidate
5 assignments for which the contents of the assigned fields
6 satisfy all of the applicable rules.

1 9. A method according to claim 1, wherein assigning the
2 labels to the rules comprises finding one or more
3 candidate assignments of all of the labels to the
4 respective fields, for which the contents of the assigned
5 fields satisfy all of the applicable rules, and
6 processing the contents of the fields responsive to the
7 one or more candidate assignments.

1 10. A method according to claim 9, wherein finding the
2 one or more candidate assignments comprises finding a
3 plurality of alternative candidate assignments for which
4 the contents of the assigned fields satisfy all of the
5 applicable rules, and comparing the candidate assignments
6 so as to unequivocally assign at least a subset of the
7 labels to the respective fields.

1 11. A method according to claim 1, wherein the document
2 comprises one of a plurality of form documents sharing a
3 common layout, and wherein assigning the labels comprises

4 making an assignment with respect to all of the form
5 documents.

1 12. A method according to claim 11, wherein making the
2 assignment comprises choosing the assignment so as to
3 satisfy a statistical criterion with respect to
4 satisfaction of the applicable rules by the contents of
5 the fields in at least a subset of the plurality of form
6 documents.

1 13. Apparatus for processing a document that includes a
2 plurality of fields having respective contents, the
3 apparatus comprising a document processor arranged to
4 receive a listing of labels to be assigned respectively
5 to the fields and one or more rules applicable to the
6 contents of the fields according to the labels assigned
7 thereto, and further arranged to read the contents of the
8 fields and to assign the labels to the fields responsive
9 to application of the rules to the contents.

1 14. Apparatus according to claim 13, wherein the rules
2 comprise a criterion to be applied to the contents of a
3 single one of the fields.

1 15. Apparatus according to claim 13, wherein the rules
2 comprise a relation between the contents of two or more
3 of the fields.

1 16. Apparatus according to claim 15, wherein the
2 contents comprise numbers, and wherein the relation
3 comprises a mathematical relationship between the numbers
4 contained in the two or more of the fields.

1 17. Apparatus according to claim 15, wherein the
2 contents comprise alphanumeric characters, and wherein

3 the relation comprises a semantic relationship between
4 words formed by the characters.

1 18. Apparatus according to claim 13, wherein the rules
2 further comprise one or more geometrical rules indicating
3 an expected geometrical relationship between two or more
4 of the fields according to the labels assigned thereto,
5 and wherein the processor is arranged to apply the
6 geometrical rules along with the rules applicable to the
7 contents of the fields so as to assign the labels to the
8 fields.

1 19. Apparatus according to claim 13, wherein the
2 processor is arranged to make a test assignment of the
3 labels to the fields, and to reject the test assignment
4 if the contents of the assigned fields do not satisfy the
5 rules applicable according to the labels.

1 20. Apparatus according to claim 19, wherein the
2 processor is arranged to iteratively test substantially
3 all possible assignments of the labels to the fields, so
4 as to find one or more candidate assignments for which
5 the contents of the assigned fields satisfy all of the
6 applicable rules.

1 21. Apparatus according to claim 13, wherein the
2 processor is arranged to find one or more candidate
3 assignments of all of the labels to the respective
4 fields, for which the contents of the assigned fields
5 satisfy all of the applicable rules, and to process the
6 contents of the fields responsive to the one or more
7 candidate assignments.

1 22. Apparatus according to claim 21, wherein the one or
2 more candidate assignments comprise a plurality of

3 alternative candidate assignments for which the contents
4 of the assigned fields satisfy all of the applicable
5 rules, and wherein the processor is arranged to compare
6 the candidate assignments so as to unequivocally assign
7 at least a subset of the labels to the respective fields.

1 23. Apparatus according to claim 13, wherein the
2 document comprises one of a plurality of form documents
3 sharing a common layout, and wherein the processor is
4 arranged to make an assignment of the labels with respect
5 to all of the form documents.

1 24. Apparatus according to claim 23, wherein the
2 processor is arranged to make the assignment so as to
3 satisfy a statistical criterion with respect to
4 satisfaction of the applicable rules by the contents of
5 the fields in at least a subset of the plurality of form
6 documents.

1 25. A computer software product for processing a
2 document that includes a plurality of fields having
3 respective contents, the product comprising a
4 computer-readable medium in which program instructions
5 are stored, which instructions, when read by a computer,
6 cause the computer to receive a listing of labels to be
7 assigned respectively to the fields and one or more rules
8 applicable to the contents of the fields according to the
9 labels assigned thereto, to read the contents of the
10 fields, and to assign the labels to the fields responsive
11 to application of the rules to the contents.

1 26. A method for computerized data processing,
2 comprising:
3 receiving information that is input to a plurality
4 of fields in a form on a computer;

5 providing labels to be assigned respectively to the
6 fields and one or more geometrical rules indicating an
7 expected geometrical relationship between two or more of
8 the fields in the form according to the labels assigned
9 to the fields; and

10 assigning the labels to the fields responsive to the
11 information and to application of the rules to the
12 fields.

1 27. A method according to claim 26, wherein the form
2 comprises a table, and wherein receiving the information
3 comprises receiving characters keyed into the fields of
4 the table by an operator.

1 28. Apparatus for data processing, comprising a form
2 processor, which is arranged to receive information that
3 is input to a plurality of fields in a form on a
4 computer, and to receive a listing of labels to be
5 assigned respectively to the fields and one or more
6 geometrical rules indicating an expected geometrical
7 relationship between two or more of the fields in the
8 form according to the labels assigned to the fields, and
9 to assign the labels to the fields responsive to the
10 information and to application of the rules to the
11 fields.

1 29. A computer software product, comprising a
2 computer-readable medium in which program instructions
3 are stored, which instructions, when read by a computer,
4 cause the computer to receive information that is input
5 to a plurality of fields in a computerized form, and to
6 receive a listing of labels to be assigned respectively
7 to the fields and one or more geometrical rules
8 indicating an expected geometrical relationship between

39943S2

9 two or more of the fields in the form according to the
10 labels assigned to the fields, and to assign the labels
11 to the fields responsive to the information and to
12 application of the rules to the fields.